

CHAPTER 3

Integration of Differential Equations

Processes which change continuously with time generally are described by one or more ordinary differential equations. For example, suppose an investment $M(t)$ draws interest continuously at a given but not necessarily constant interest rate $r(t)$. At the same time money is added or withdrawn at a prescribed continuous rate $b(t)$. Clearly one would like to know $M(t)$ over the life $[0, T]$ of the contract.

The mathematical model for this process is a simple balance equation:

rate of change of the investment = money coming in – money going out,

or

$$\frac{dM(t)}{dt} = r(t)M(t) + b(t).$$

We also know $M(t)$ either initially or at the end of the contract. To be specific let us assume that $M(t_0) = M_0$ is given.

A more suggestive way of writing this equation is

$$LM(t) \equiv \frac{dM}{dt} - r(t)M = b(t).$$

L is a linear differential operator since $L[x(t) + \alpha y(t)] = Lx(t) + \alpha Ly(t)$ for arbitrary differentiable functions $x(t)$ and $y(t)$. L is called a first order differential operator because only the first derivative appears, it has variable coefficients since $r(t)$ is not assumed to be constant, and the equation $LM = b$ is inhomogeneous since $M = 0$ is not a solution. Accordingly, the differential equation is known as a first order linear inhomogeneous differential equation with variable coefficients.

Such an equation can be solved analytically. It follows from the general theory that its solution is of the form

$$M(t) = M_c(t) + M_p(t)$$

where $M_c(t)$ is known as the complementary solution which solves

$$LM_c(t) = 0$$

and which includes the constant of integration, and where $M_p(t)$ is a particular integral which means ANY function which solves

$$LM_p(t) = b(t).$$

Once $M_c(t)$ and $M_p(t)$ have been found the constant of integration is computed such that $M(t_0) = M_0$.

The complementary solution is found by the method of separation of variables. We rewrite the equation

$$M'(t) - r(t)M(t) = 0$$

in the form

$$\frac{dM}{M} = r(t)dt$$

and integrate the left side with respect to M and the right side with respect to t .

As an aside we point out that the integration of one side with respect to M and the other side with respect to t is not at all strange. In general, when we integrate a separable equation like

$$f(M)dM = g(t)dt$$

we obtain an expression like

$$F(M) = G(t) + k$$

where F and G are the anti-derivatives of f and g . This equation defines M implicitly in terms of t . Implicit differentiation and the chain rule then show that

$$\frac{d}{dt} F(M(t)) = \frac{dF}{dM} \frac{dM}{dt} = f(M)M' = \frac{dG}{dt} = g(t)$$

so that $M(t)$ satisfies the separable differentiable equation. In our problem the integration yields

$$\log |M| = \int_{t_0}^t r(x)dx + k$$

where k is a constant of integration. The lower limit on the integral was chosen for convenience. Changing it is equivalent to changing k . It follows that

$$M_c(t) = ce^{\int_{t_0}^t r(x)dx}$$

for any non-zero constant c .

The particular integral can be computed with the method of variation of parameters. We are looking for a solution of the form

$$M_p(t) = M_c(t)v(t)$$

where $v(t)$ is to be found so that $M_p(t)$ satisfies the inhomogeneous equation. We find

$$M_p'(t) = M_c'(t)v(t) + M_c(t)v'(t) = r(t)M_p(t) + b(t).$$

Since $M_c' = r(t)M_c$ the two terms involving $r(t)$ cancel and we are left with

$$M_c(t)v'(t) = b(t)$$

so that

$$v(t) = \int_{t_0}^t M_c^{-1}(s)b(s)ds.$$

Again, the lower limit is chosen for convenience. Any fixed lower limit could be used. Since

$$M_c^{-1}(s) = \frac{1}{c}e^{-\int_{t_0}^s r(x)ds}$$

we find that

$$M_p(t) = \int_{t_0}^t e^{\int_s^t r(x)ds} b(s)ds.$$

The final solution then is

$$M(t) = M_0e^{\int_{t_0}^t r(x)dx} + M_p(t)$$

which is seen to satisfy $M(t_0) = M_0$. This formula was developed under the assumptions that $r(t)$ and $b(t)$ are continuous functions of t . However, it is readily modified when $r(t)$ is only piecewise continuous and $b(t)$ may in fact describe a single payment at a specified time. For example, since only the integral of r appears in our solution we need just integrability of r . Hence finitely many jumps do not cause any problems. The source term $b(t)$ needs a little more care if we are to model something like a fixed coupon payment of known size at time t_0 . To see how the above solution can be modified let us think of a bond which pays $\$M_0$ at time t_0 and a single coupon at time $t_c < t_0$ in the amount of $\$C$. Since the

bond must decrease in value by the amount of the coupon at the time of its payment we have a withdrawal from the account. We shall employ the usual notation of $t_0 = T$ for the time when the bond matures. Were the coupon payment stretched out over an interval and paid at a rate $b(t)$ then the value $M(t)$ of the bond is given by the above expression after reversing the direction of integration as

$$M(t) = M_0 e^{-\int_t^T r(x)dx} + \int_t^T e^{-\int_t^s r(x)dx} b(s)ds$$

where the algebraic signs are chosen such that $b(t) > 0$. Let us approximate the one time payment by a time continuous payment of rate $\frac{C}{\Delta t} \phi_c(t)$ where $\phi_c(t)$ is the piecewise linear hat function with values $\phi_c(t_c - \Delta t) = \phi_c(t_c + \Delta t) = 0$ and $\phi_c(t_c) = 1$. Then $\frac{C}{\Delta t} \int_a^b \phi_c(t)dt = C$ if the interval $[a, b]$ includes the interval $[t_c - \Delta t, t_c + \Delta t]$ because the area under the hat function is $1/\Delta t$. Let us now suppose that $t > t_c + \Delta t$. Then $b(s) = 0$ for all $s \in [t, T]$. Hence the particular integral is zero and

$$M(t) = M_0 e^{-\int_t^T r(x)dx}.$$

This is the usual discounted value of a zero coupon bond which pays M_0 at maturity. Let us suppose next that $t < t_c - \Delta t$. Then $b(s)$ is non-zero only for $s \in [t_c - \Delta t, t_c + \Delta t]$ and the particular integral becomes

$$M_p(t) = \frac{C}{\Delta t} \int_{t_c - \Delta t}^{t_c + \Delta t} e^{-\int_t^s r(x)ds} \phi_c(s)ds.$$

The mean value theorem for integrable functions applies because $\phi_c(t)$ is non-negative. It yields

$$M_p(t) = \frac{C}{\Delta t} e^{-\int_t^{\hat{s}} r(x)dx} \int_{t_c - \Delta t}^{t_c + \Delta t} \phi_c(s)ds = C e^{-\int_t^{\hat{s}} r(x)dx}$$

where \hat{s} is some point in the interval $[t_c - \Delta t, t_c + \Delta t]$. A coupon payment at the fixed time t_c is obtained if we let $\Delta t \rightarrow 0$. This means that $\hat{s} \rightarrow t_c$. Hence the value of our coupon bearing bond is

$$M(t) = M_0 e^{-\int_t^T r(x)dx}, \quad t > t_c$$

$$M(t) = M_0 e^{-\int_t^T r(x)dx} + C e^{-\int_t^{t_c} r(x)dx}, \quad t < t_c.$$

The last integral is just the discounted value of the coupon paid at time t_c . It is apparent that the “analytic” solution provides useful qualitative insight even when the integrals in the formula cannot be evaluated in closed form. However, quantitative data are obtainable only if all integrations can be carried out. This is usually not the case unless the interest rate is (piecewise) constant and b is sufficiently simple. Otherwise one is forced into the numerical evaluation of the integrals, but then one may as well integrate the differential equation itself numerically.

Numerical methods for ordinary differential equations apply equally to linear and non-linear equations, to a single equation or to a system of equations. This generality is essential because such problems are commonplace. For example, take a simple model of an investment where the interest rate depends on the amount invested such as

$$\frac{dM}{dt} = r(M)M + b(t)$$

$$M(t_0) = M_0$$

where

$$r(M) = \begin{cases} r_u & \text{for } M > M_u \\ r_1 \frac{M_u - M}{M_u - M_1} + r_u \frac{M - M_1}{M_u - M_1} & M_1 < M < M_u \\ r_1 & \text{for } M < M_1 \end{cases}$$

where the thresholds and rates may in fact be time dependent. For a numerical integration this problem is not much more complicated than the simple model solved analytically above (although the lack of differentiability of the right side may require a little care before applying standard integration programs). Similarly, suppose we have N accounts and moneys flow between them in proportion to their size. A model might look like

$$\frac{dM}{dt} = A(M, t)M + b(t)$$

where now $M(t) = \begin{pmatrix} M_1(t) \\ \dots \\ M_N(t) \end{pmatrix}$ and A is an $N \times N$ matrix whose entries could depend on M .

Again, numerical methods would not be troubled by such a system of differential equations. We remark that if A does not depend on M then one can find a vector analog of the above analytic solution, but its numerical evaluation is problematic and generally not easier than a numerical integration of the differential equation.

Numerical methods for ordinary differential equations have reached a high level of perfection and one has generally as little difficulty solving numerically a problem like

$$u' = F(t, u)$$

$$u(t_0) = u_0$$

with standard library programs as one would have solving the linear algebraic system

$$Ax = b.$$

We do not intend to study the various approaches to solving a first order system. For well behaved systems whose solutions change smoothly at comparable rates any of the standard routines will do a good job. The only advice is to use one of the highly developed library routines and not to try to code your own numerical method. Only if the system should resist a numerical integration with your chosen solver does a differentiation among the various approaches appear justified. But then deeper insight is required than we can provide at this point.

Nonetheless, so that the numerical integration of a differential equation is not entirely relegated to a black box we shall discuss in some detail the simplest numerical method in the simplest setting of just one equation. This discussion will give us the flavor of what numerical methods are based on and will allow the transition to simulating the solution of a stochastic ordinary differential equation.

Euler's method for ordinary differential equations: Let $u(t)$ be the solution of

$$u' = F(t, u)$$

$$u(t_0) = u_0$$

then at t_0 we have the solution u_0 and the slope $u'(t_0) = F(u_0, t_0)$. We can approximate the solution $u(t_0 + \Delta t)$ for some (small) Δt by the value of the linear approximation to $u(t)$ at $t_1 = t_0 + \Delta t$, i.e.

$$u_1 = u_0 + \Delta t F(t_0, u_0).$$

Given (t_1, u_1) we then find the straight line approximation at $t_2 = t_1 + \Delta t$ with slope $F(t_1, u_1)$. Continuing in this way a piecewise linear curve is generated through the points $\{(t_n, u_n)\}$ where

$$t_{n+1} = t_n + \Delta t$$

and

$$u_{n+1} = u_n + \Delta t F(t_n, u_n).$$

The obvious question now is how well the computed values $\{u_n\}$ approximate the values of the true solution $\{u(t_n)\}$. The following result applies:

Theorem: *Suppose that the problem*

$$u' = F(t, u)$$

$$u(t_0) = u_0$$

has a solution $u(t)$ over the interval $[t_0, T]$ which satisfies

$$|u''(t)| < K \quad \text{for all } t \in [t_0, T].$$

Suppose further that

$$\left| \frac{\partial F}{\partial u}(t, u) \right| \leq L$$

for all u and for all $t \in [t_0, T]$. Let $\Delta t = (T - t_0)/N$, $t_n = t_0 + n\Delta t$ and let $\{u_n\}$ denote the values computed with Euler's method. Define the error of the approximation as

$$e_n = u(t_n) - u_n$$

then

$$|e_n| < e^{L(t_n - t_0)} |e_0| + \frac{K\Delta t}{2L} \left(e^{L(t_n - t_0)} - 1 \right).$$

Proof: It follows from a Taylor expansion that the analytic solution satisfies

$$u(t_{n+1}) = u(t_n) + u'(t_n)\Delta t + \frac{1}{2} u''(\eta)\Delta t^2,$$

i.e.,

$$u(t_{n+1}) = u(t_n) + F(t_n, u(t_n))\Delta t + \frac{1}{2} u''(\eta)\Delta t^2$$

for some $\eta \in (t_n, t_{n+1})$. Subtracting the Euler approximation we obtain

$$e_{n+1} = e_n + [F(t_n, u(t_n)) - F(t_n, u_n)]\Delta t + \frac{1}{2} u''(\eta)\Delta t^2.$$

The mean value theorem for differentiable functions yields

$$|F(t_n, u(t_n)) - F(t_n, u_n)| = \left| \frac{\partial F}{\partial u}(t_n, \xi)(u(t_n) - u_n) \right|$$

for some ξ between $u(t_n)$ and u_n . Since the partial derivative is bounded we obtain

$$|F(t_n, u(t_n)) - F(t_n, u_n)| \leq L|u(t_n) - u_n| = L|e_n|.$$

We can estimate the error at t_{n+1} according to

$$|e_{n+1}| < (1 + L\Delta t)|e_n| + \frac{\Delta t^2}{2} K.$$

By induction (or repeatedly substituting) we see that this inequality implies

$$|e_n| < (1 + L\Delta t)^n |e_0| + \frac{\Delta t^2}{2} K \sum_{j=0}^{n-1} (1 + L\Delta t)^j.$$

A Taylor expansion about $x = 0$ shows that for any x

$$e^x = 1 + x + \frac{e^\xi}{2} x^2$$

for some ξ between x and 0. Hence

$$1 + x \leq e^x$$

which implies that

$$(1 + L\Delta t)^n < e^{Ln\Delta t} = e^{L(t_n - t_0)}.$$

The sum in the error bound is a geometric progression and can be evaluated analytically. If we have

$$S_N = \sum_{j=0}^N r^j$$

then we obtain for $r \neq 1$

$$(1 - r)S_N = (1 - r) \sum_{j=0}^N r^j = 1 - r^{N+1}$$

so that

$$S_N = \frac{1 - r^{n+1}}{1 - r}.$$

(Note that by l'Hospital's rule

$$\lim_{r \rightarrow 1} S_N = N + 1.)$$

This summation formula implies that

$$\sum_{j=0}^n (1 + L\Delta t)^j = \frac{(1 + L\Delta t)^{n+1} - 1}{L\Delta t} \leq \frac{e^{L(t_n - t_0)} - 1}{L\Delta t}.$$

When we use these estimates in our error formula we obtain the final result

$$|e_n| < e^{L(t_n - t_0)} |e_0| + \frac{K\Delta t}{2L} (e^{L(t_n - t_0)} - 1).$$

What does this error bound tell us about the quality of the numerical solution? The first observation is that we must expect the error to increase as we take more and more time steps of constant size since the exponential term will grow. On the other hand, if we integrate the equation over a fixed time interval $[t_0, T]$ then the exponential is bounded by $e^{L(T - t_0)}$ so that the error depends on e_0 and Δt . The appearance of e_0 might be puzzling because we set u_0 equal to the given initial condition. However, the given initial condition for the differential equation may not be one of the numbers representable in the computer, which implies that e_0 is not in general zero but proportional to the accuracy with which the computer approximates a given number. On modern long wordlength machines this number tends to be of order 10^{-7} and is negligible. The second error term dominates. It goes to zero like Δt . The common terminology which expresses this behavior is that Euler's method converges with order Δt . This rate is a very slow compared to the commonly used integrators of program libraries, such as Runge-Kutta, Runge-Kutta-Fehlberg or Adams integrators whose convergence is of order Δt^4 or Δt^5 .

We shall conclude this brief introduction into the numerical solution of ordinary differential equations with an example of a reasonable looking numerical method but which nonetheless can give nonsensical results for a simple model problem. This will be our first exposure to an unstable numerical method and foreshadows some of the problems we need

to contend with when we integrate the Black Scholes pricing equation in Chapter 4. We shall consider the problem

$$\begin{aligned}u' &= -2u + 1 \\u(0) &= 1.\end{aligned}$$

The analytic solution is

$$u(t) = \frac{1}{2} (e^{-2t} + 1).$$

Suppose that we are interested in the numerical solution over the interval $[0, T]$. A numerical approximation to the differential equation is the finite difference equation

$$\frac{u_{n+1} - u_{n-1}}{2\Delta t} = -2u_n + 1$$

where u_n is the approximation to $u(t_n)$ at $t_n = n\Delta t$ with $\Delta t = T/N$ for a given $N > 0$. This formula is not self-starting since we need u_0 and u_1 in order to compute u_n for $n = 2, 3, \dots, N$. However, we can always find a good approximation to $u(t_1)$, and hence for u_1 , from the Taylor expansion

$$u(t_n) = \sum_{k=0}^M \frac{u^{(k)}(0)}{k!} \Delta t^k$$

because u and its derivatives are available recursively from the initial value u_0 and the differential equation. To be specific we have

$$\begin{aligned}u(0) &= 1 \\u'(0) &= -2u(0) + 1 \\u''(0) &= -2u'(0) \\&\dots\dots\dots \\u^{(M)}(0) &= -2u^{(M-1)}(0).\end{aligned}$$

With u_0 and u_1 given we have to solve for $n = 2, \dots, N$ the difference equation

$$u_{n+1} + 4\Delta t u_n - u_{n-1} = 2\Delta t.$$

It is known that the solution to such a difference equation can be expressed in form

$$\vec{u} = \vec{u}_c + \vec{u}_p$$

where $\vec{u}_c = (u_0, \dots, u_N)$ is a solution of the homogeneous difference equation

$$u_{n+1} + 4\Delta t u_n - u_{n-1} = 0$$

and \vec{u}_p is ANY vector which satisfies

$$u_{n+1} + 4\Delta t u_n - u_{n-1} = 2\Delta t.$$

It is straightforward to verify that a constant vector u_p with every component u_n given by

$$u_n = \frac{1}{2}$$

solves the inhomogeneous equation. It remains to find a general \vec{u}_c which will allow us to incorporate the initial conditions. We look for a solution of the form

$$u_n = \beta^n \quad \text{with} \quad \beta \neq 0.$$

Substitution into the homogeneous equation shows that β must be a solution of the quadratic equation

$$\beta^2 + 4\Delta t \beta - 1 = 0.$$

There are two roots

$$\beta_1 = -2\Delta t + \sqrt{4\Delta t^2 + 1}$$

$$\beta_2 = -2\Delta t - \sqrt{4\Delta t^2 + 1}.$$

The n th component of our general solution can now be written as

$$u_n = c_1 \beta_1^n + c_2 \beta_2^n + \frac{1}{2}$$

where c_1 and c_2 are computed such that u_0 and u_1 take on the prescribed initial conditions.

Let us now look at the value u_N as $N \rightarrow \infty$, i.e., as $\Delta t \rightarrow 0$. We write

$$\beta_1^N = e^{N \log \beta_1}$$

and compute

$$\lim_{N \rightarrow \infty} N \log \left(-\frac{2T}{N} + \sqrt{\frac{4T^2}{N^2} + 1} \right).$$

Since this yields formally

$$0 \cdot \infty$$

we apply l'Hospital's rule to

$$\lim_{N \rightarrow \infty} \frac{\log \left(-\frac{2T}{N} + \sqrt{\frac{4T^2}{N^2} + 1} \right)}{\frac{1}{N}}$$

and obtain

$$\lim_{N \rightarrow \infty} \frac{\log \left(-\frac{2T}{N} + \sqrt{\frac{4T^2}{N^2} + 1} \right)}{\frac{1}{N}} = -2T$$

so that

$$\lim_{N \rightarrow \infty} \beta_1^N = e^{-2T}.$$

An analogous argument shows that

$$\lim_{N \rightarrow \infty} |\beta_2^N| = e^{2T}.$$

Hence the startling conclusion is that

$$u_N = c_1 \beta_1^N + c_2 (-1)^N |\beta_2|^N + \frac{1}{2}$$

where the term involving β_1 converges to the correct analytic solution while the other term oscillates and converges to $c_2 e^{2T}$ in magnitude. Clearly for large T this term will dominate the other two terms so that the numerical solution can have a meaning only if $c_2 = 0$. However, since a computer does not perform exact arithmetic the exponentially growing term cannot be suppressed. The above finite difference approximation is said to be unstable because a small perturbation will grow and destroy the solution.

As an illustration we show in Fig. 1a a plot of the analytic, the Euler and the above approximate solution over $[0, 1]$ for $\Delta t = 1/50$ when u_1 is equal to the analytic solution $u(t_1)$. The positive exponential is small in this case and effectively suppressed by the difference method. However, if we integrate the equation over $[0, 5]$ with the same Δt then the analytic, Euler and central difference solutions shown in Fig. 1b result. The oscillating positive exponential term dominates for larger t and destroys the meaning of the numerical solution.

The oscillations can be suppressed over a larger interval by taking a smaller Δt (which is reflected, theoretically, in a smaller c_2), but for a sufficiently large T they will eventually take over. It is important to recognize that some numerical methods are inherently unstable and therefore to be avoided. Large unexpected oscillations in a computed solution are usually a clue that one should look for another numerical method.

Let us conclude this discussion of our numerical method by introducing the concept of extrapolation which is valuable not only in the numerical solution of differential equations, but also in the context of numerical integration and, in fact, anytime where a numerical process depends on a (small) parameter h such that the problem has a true solution u and the algorithm provides a numerical solution $U(h)$ for which one has an estimate of the form

$$u - U(h) = Kh^\alpha + O(h^\beta)$$

where $O(h^\beta)$ denotes a term which is negligible compared to the term Kh^α . For example, the proof of the convergence theorem for Euler's method with $h = \Delta t$ showed that

$$u - U(h) = Kh + O(h^2)$$

where u and $U(h)$ denote the analytic and numerical solution at an arbitrary point t_n .

Let us suppose we know the exponent α and have computed numerical values for $U(h)$ and $U(h/2)$. Then it follows from

$$u - U(h) = Kh^\alpha + O(h^\beta)$$

and

$$u - U(h/2) = K(h/2)^\alpha + O(h^\beta)$$

that

$$u - U(h) - 2^\alpha(u - U(h/2)) = O(h^\beta) - 2^\alpha O(h^\beta).$$

If we solve for u we find

$$(2^\alpha - 1)u = 2^\alpha U(h/2) - U(h) + O(h^\beta)$$

so that

$$U = \frac{2^\alpha U(h/2) - U(h)}{2^\alpha - 1}$$

is a better approximation to u than either $U(h)$ or $U(h/2)$ because the dominant error term has been eliminated. U is considered an extrapolation of the values $U(h/2)$ and $U(h)$. But note that α has to be known in order to find U .

Now given a numerical method depending on a parameter h one may suspect that its approximate solution satisfies the relation

$$u - U(h) = Kh^\alpha + O(h^\beta)$$

but short of a complete analysis analogous to our convergence proof for Euler's method we may not know the exponent α . If the relationship is indeed true one can actually determine it experimentally. Thus, suppose we have three numerical results $U(h)$, $U(h/2)$ and $U(h/4)$. Then by hypothesis

$$\begin{aligned}u - U(h) &= Kh^\alpha + O(h^\beta) \\u - U(h/2) &= K(h/2)^\alpha + O((h/2)^\beta) \\u - U(h/4) &= K(h/4)^\alpha + O((h/4)^\beta)\end{aligned}$$

where u , K and α are not known. Since the remainder term $O(h^\beta)$ is supposed to be small compared to the dominant error term Kh^α we shall ignore it henceforth. By subtracting one relationship from another we obtain

$$U(h/2) - U(h) = K(h^\alpha - (h/2)^\alpha)$$

and

$$U(h/4) - U(h/2) = K((h/2)^\alpha - (h/4)^\alpha).$$

Division now yields

$$\frac{U(h/2) - U(h)}{U(h/4) - U(h/2)} = \frac{(1 - 1/2^\alpha)}{(1/2^\alpha - 1/4^\alpha)} = 2^\alpha.$$

Since numbers are available from the computation for the left side we can solve for α . Once α is known we can use the data $U(h/2)$ and $U(h/4)$ to extrapolate an improved approximation

$$U = \frac{2^\alpha U(h/4) - U(h/2)}{2^\alpha - 1}.$$

To illustrate this process and to give some feel for the numbers one might expect, let us look at the following example.

An account with an initial investment of $M(0) = \$2$ (say in millions) draws continuously compounded interest at a constant rate of 8%. At the same time we pay continuously the tax on the account which is assessed at a varying rate between 1.807% and 2.2% according to

$$.02(1 + 0.1 \tanh(2(M(t) - 1))).$$

The differential equation for the money $M(t)$ in the account is

$$M'(t) = .08M(t) - [.02(1 + 0.1 \tanh(2(M(t) - 1)))]M(t)$$

$$M(0) = 2$$

where t is measured in years. This equation will be integrated with Euler's method for a time span of four years. We show in Table 1 the number of time steps, the time step Δt , the numerical value of $M(4)$ obtained with Euler's method, the estimate for the exponent α for the order of convergence based on the three runs with time step Δt , $2\Delta t$ and $4\Delta t$, and finally the extrapolated value of $M(4)$ computed from the values of $M(4)$ for the time steps Δt and $2\Delta t$ and the theoretical order of convergence of $\alpha = 1$.

Table 1: A record of Euler's method for the model problem

N	T	$M(4)$		extrapolated $M(4)$
20	4/20	2.51920		
40	4/40	2.52087		2.52254
80	4/80	2.52171	.9909	2.52255
160	4/160	2.52214	.9955	2.52256
320	4/320	2.52235	.9943	2.52256
640	4/640	2.52245	.9919	2.52256
1280	4/1280	2.52250	1.0866	2.52255
2560	4/2560	2.52253	.8074	2.52256
5120	4/5120	2.52254	1.2630	2.52256

These numbers are typical for what one would expect. The estimate for the order of convergence is near unity but degrades noticeably as Δt becomes small. This is due to the division

of one very small number by another and suggests that this calculation should be carried out in double precision. The remarkable result is that the extrapolation of the results from even the two coarsest grids already yields the numerical solution to which Euler's method converges as the grids become finer and finer.

However, it must again be pointed out that Euler's method (even with extrapolation) is a poor choice for the numerical integration of initial value problems for ordinary differential equations. Higher order Runge-Kutta and Adams methods should be employed.

Stochastic differential equations: When one prices an option with a maturity of several days or weeks in the future then must have a view of how the value of the underlying asset will change over this period. Of course, this value cannot be predicted with any certainty. In fact, its short term evolution looks quite random. Such motion is modeled with a stochastic differential equation which describes a so-called Itô process:

$$dx = a(t, x)dt + b(t, x)dW$$

$$x(t_0) = x_0$$

where a and b are given functions of t and x .

This equation is far removed from the ordinary differential equation discussed above. In this setting x is a random variable because $W(t)$ is the random variable associated with a standard Brownian motion which is defined as follows []:

Definition: A scalar standard Brownian motion, or standard Wiener process, over $[0, T]$ is a random variable $W(t)$ that depends continuously on $t \in [0, T]$ and satisfies the following three conditions:

- 1) $W(0) = 0$ (with probability 1).
- 2) For $0 \leq s < t \leq T$ the random variable given by the increment $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$: equivalently $W(t) - W(s) \in \sqrt{t - s} N(0, 1)$, where $N(0, 1)$ denotes a normally distributed random variable with zero mean and unit variance.
- 3) For $0 \leq s < t < u < v \leq T$ the increments $W(t) - W(s)$ and $W(v) - W(u)$ are independent.

Our goal will be to integrate numerically the stochastic differential equation. However, we cannot compare numerical results with anything analytic unless we have the same path for the Brownian motion. Hence we shall fix the sample path with a discrete Brownian motion.

Discretized Brownian motion: We set $\delta t = T/N$. W_n is the numerical value of $W(t)$ at $t_n = n\delta t$ obtained from $W_0 = 0$ and

$$W_n = W_{n-1} + dW_n$$

where each dW_n is an independent random variable of the form $\sqrt{\delta t} N(0, 1)$.

Let us now turn to the stochastic differential equation. The equation itself is just a shorthand notation for the integral equation

$$x(t) = x_0 + \int_{t_0}^t a(s, x(s))ds + \int_{t_0}^t b(s, x(s))dW(s).$$

We shall assume that the functions a and b are well behaved deterministic functions so that the first integral is unambiguous. The second integral, on the other hand, is going to be interpreted as a so-called Itô integral

$$\int_0^T b(s, x(s))dW(s) = \lim_{\Delta t \rightarrow 0} \sum_{j=0}^{K-1} a(t_j, x(t_j))(W(t_{j+1}) - W(t_j))$$

where $\Delta t = T/K$ and $t_{j+1} - t_j = \Delta t$. It is now natural to integrate the differential equation over $[0, T]$ with Euler's method

$$x_{j+1} = x_j + a(t_j, x_j)\Delta t + b(t_j, x_j)(W(t_{j+1}) - W(t_j))$$

where x_0 is the given initial value and Δt is a fixed (constant) step size. This formula is unambiguous. Once we have found x_j we draw the random number $W(t_{j+1}) - W(t_j)$ from $\sqrt{\Delta t} N(0, 1)$ and compute x_{j+1} . The points $\{x_j\}$ form a sample path. If we repeat the calculation an entirely different sample path will be found. If we find M sample paths we can compute the mean of $\{x_j\}$ and the deviation from the mean for all the sample paths which provides statistical information on the solution of the stochastic differential equation.

As in the deterministic case one can discuss the error between the analytic and the numerical solution. This can make sense only if both solutions are based on the same sample path for the Brownian motion. This is where the discrete Brownian motion comes into play. We shall assume that the points $\{t_j\}$ used for Euler's method form a subset of the points $\{t_n\}$ where we use the values $\{W(t_j)\}$ given by the discrete Brownian motion. Even though the sample path for the continuous solution and the Euler solution have exactly the same random behavior, the two solutions will differ because the integrals are approximated by finite sums. If one were to know the exact solution of the stochastic differential equation then, of course, one can compare the influence of the step size Δt on the accuracy of the numerical solution with respect to the underlying discrete Brownian motion. However, if one chooses a different sample path for the discrete Brownian motion then a new behavior of the error would be observed. This raises the question of whether one can say something about the convergence of the numerical solution independently of the underlying sample path. The answer is to measure the expected value of the error obtained for many different sample paths.

Definition: Let $x(t_j)$ and x_j be the random variables which solve the stochastic differential equation and a discrete approximation to it obtained for a step size of Δt . Then the numerical method has a STRONG order of convergence γ if there exists a constant C such that the expectation of the error satisfies

$$E|x_j - x(t_j)| \leq C\Delta t^\gamma.$$

Here C will depend on the solution $x(t)$ and the interval of integration $[0, t_j]$ but not on Δt . It is known that for smooth functions a and b the above Euler method has strong order of convergence given by

$$\gamma = 1/2.$$

We note that numerical experiments are reported in [] for the stochastic differential equation

$$dx = \lambda x + \mu x dW$$

$$x(0) = x_0$$

which has the analytic solution

$$x(t) = x(0) \exp((\lambda - \mu^2/2)t + \mu W(t)).$$

Thousand different discrete Brownian motion sample paths are generated over $[0, 1]$ for $\lambda = 2$, $\mu = 1$, $x_0 = 1$ and $\delta t = 2^{-8}$. A thousand errors $|x(T) - x_K|$ are averaged for each of five different step sizes Δt . The logarithm of the sample averages are plotted against the logarithm of Δt . If the strong order of convergence inequality is in fact roughly an equality then

$$\log E|x(T) - x_K| = \log C + \gamma \log \Delta t,$$

i.e. a linear relationship would be expected. The five data points for the five different step sizes do not lie exactly on a straight line, but a least squares straight line approximation to the data exhibits a slope of $\gamma = .5383$ which is taken as consistent with the prediction of $\gamma = .5$.

Since any one particular sample path $x(t)$ in itself has little meaning it is sensible to compare also the mean of numerical and analytic solutions. This leads to the concept of a weak order of convergence.

Definition: A numerical method has weak order of convergence equal to γ if there exists a constant C such that for all functions g (belonging to a general specified class)

$$|E(g(x(t_j)) - Eg(x_j))| \leq C\Delta t^\gamma$$

where C generally will depend on g , on $x(t)$ and on t but not on Δt . It can be shown that for Euler's method the weak order of convergence is equal to 1. This implies that means and variances are approximated to order Δt with the numerical method.

In finance the most common stochastic differential equations are the model for the value S of an equity asset

$$dS = r(S, t)S dt + \sigma(S, t)S dW$$

where $r(S, t)$ represents an interest term and $\sigma(S, t)$ is the volatility of S , and the model for a short term interest rate like

$$dr = a(b - r)dt + \sigma dW.$$

While it may be instructive to compute sample paths for these equations and to compute means and variances, their primary use will be deriving pricing equations for options and bonds which depend on the random functions S and r , respectively.